

Unknown Value Reports

Frequently Asked Questions

Note: We use the following conventions in this document: **Bold** denotes a table name, while *italics* denote a field name.

1. Q: Why does DMHAS produce Unknown Values Reports?
A: The DMHAS Unknown Values Reports serve two purposes:
 - Display in a concise format the amount of missing data in the DMHAS database for each provider and treatment program.
 - Enable DMHAS and providers to cross-check their counts of admissions, discharges, unduplicated consumers, etc., in order to track down any discrepancies between providers' and DMHAS' data sources.
2. Q: What do the Unknown Values Reports cover?
A: The Unknown Values Reports cover fields included in the DPAS Data Dictionary, which is available on-line at <http://www.dmhas.state.ct.us/QMI/datacollection.pdf>. The Reports cover all programs that are:
 - Funded by DMHAS
 - Marked as 'Active' in the database
3. Q: What do the Unknown Values Reports not cover?
A: The Unknown Values Reports do not display detail on exactly which records contained Unknown values. We are working out how best to get this private medical information to you securely.
4. Q: How can I use these reports?
A: These reports can help you to determine fields in which you are consistently missing data. You can determine if certain programs are consistently missing data in certain areas. The reports show the most recent data for all months displayed, so you can track trends in data quality, as well as use prior reports to determine the effect that your data cleanup operations have had on your data quality.
5. Q: What can I do from my end to help increase data quality?
A: There are a number of DPAS and WebSAS reports that can help you look at your data in different ways.

If you submit your DPAS data via an Extract file, then you can check your file layout against the Crosswalk defined for importing your data into DPAS. This Crosswalk is accessible via the CC872 report in the DPAS application. Many data quality problems can be cleaned up by ensuring that your Extract data is being translated correctly into DPAS.
6. Q: What is the submission deadline for data to be included on the report?
A: To be included on the report, data must be submitted to DMHAS by midnight at the end of the 15th calendar day of the month. If the 15th is a weekend or a holiday, then the data must be submitted by midnight at the end of the next business day.

7. Q: Please explain the shading that appears on the report.
A: Field names that are shaded in gray belong to the federal Treatment Episode Dataset (TEDS); these fields must be reported regularly to the federal government. The yellow shading indicates fields where the percentage of missing data for the reporting period exceeds a specified threshold. For fields not belonging to the TEDS dataset, this threshold is 10% missing data. For fields in the TEDS dataset, this threshold is 5% missing data.
8. Q: The report displays at least three months' worth of results. When I fix data from a prior month, will my changes show up on later reports?
A: Yes.
9. Q: What constitutes an Unknown value on the report?
A: An Unknown value is one of the following:
- A literal value of 'Unspecified' or 'Not Specified' in the database
 - A field in the database that is left blank
 - A field in the database containing a code not recognized in the Data Dictionary.
10. Q: How does the report calculate the number of Programs that reported data for a report period?
A: The report calculates the total number of Programs that should report data by counting the number of Programs that meet the following criteria:
- Are marked as 'Active' in the database
 - Are funded by DMHAS

For each Program, the report logic checks whether any of the following have occurred:

- The Program reported at least one admission, discharge, SATIS admission assessment, or SATIS discharge assessment during the report period.
- There was at least one claim record associated with the program having either a Service Start Date or a Service End Date during the report period.

11. Q: The DPAS table layout is different from how our database stores data. Can you give me an overview of the DPAS layout?
A: **Client** is the base table from which all other treatment data derives. Each row in the **Admission_History** table represents a single admission to a treatment program, which is identified by the *Pgm ID* field. Admissions are linked to the **Client** table using a *Client ID* field.

The **SATIS_Adm** and **SATIS_Dis** tables contain SATIS assessment data collected at admission and discharge only for consumers receiving substance abuse treatment. These tables are linked back to the **Admission_History** table using an ID number (*Admit ID*) assigned to each admission. Since SATIS records can be linked to admissions, they can in turn be linked to providers and their programs.

The **Demographic_History** table, which stores consumer Address and Marital Status information, and the **Insurance_History** table link directly to the Client table using a *Client ID* field. The Date From fields in these tables specify on which date the information in each row became effective. These tables are not linked to specific admissions, nor do these tables specify which provider entered the data. The major effect of this is that records from these tables cannot be directly connected to providers. The report handles this scenario by displaying the **Demographic_History** and **Insurance_History** data for all consumers that a) the provider admitted during the report period, and b) had a Date From within the report period.

In contrast, the **Living_Situation** table contains a program ID number (*Pgm ID*). This field enables records in this table to be linked to programs, which in turn allows them to be linked to the provider that entered the data. For this reason, **Living_Situation** data appears on the program-level reports.

12. Q: Why are the denominators for fields in the **Demographic_History** table greater than the denominators in fields for the **Client** table?

A: Multiple **Demographic_History** records can exist for a single **Client** record. A new **Demographic_History** record is created a) whenever a new consumer is entered into the database, or b) whenever an existing consumer's City or ZIP Code changes.

Each **Demographic_History** record is automatically given an *Effective Date*. When Unknown Value Reports are run for a specified period, they check all of the **Demographic_History** records that have *Effective Dates* within the period. Even if a consumer was not admitted into treatment during the report period, if he/she has **Demographic_History** records with effective dates within the period, then these records will be included in the report for that period.

A similar methodology is used to capture records in the **Insurance_History** and **Living_Situation** tables.

13. Q: How are the denominators for the different tables calculated?

A: **Client** table: The denominator is the number of records in the **Client** table that have an admission during the report period. For example, if in July 2006 a provider admitted 120 distinct consumers, then the denominator for all fields in the **Client** table will be 120.

Admission_History table: The **Admission_History** table contains both *Admission* and *Discharge Dates*, as well as Consumers' *Discharge Status*.

- For the *Admission Type* field, the denominator is simply the number of new admissions during the report period.
- For the *Discharge Type* field, the denominator is the number of records in the **Admission_History** table that contain a *Discharge Date* that falls within the report period.
- Therefore, if a program has 10 admissions and 5 discharges during July 2006, then the Unknown Values Report for that month will evaluate the *Admission Type* field for the 10 admissions and will evaluate the *Discharge Type* field for the 5 **Admission_History** records having *Discharge Dates* within July.

Demographic_History table: Multiple **Demographic_History** records can appear for each consumer.

- Each **Demographic_History** record contains an *Effective Date*.
- A new record is created for a consumer whenever the consumer's City or ZIP Code changes.
- For each reporting period, the Unknown Values Report logic captures all **Demographic_History** records having *Effective Dates* within the appropriate period.

Insurance_History table:

- Multiple **Insurance_History** records can appear for each consumer. Records in the **Insurance_History** table contain an *Effective Date*.
- A new record is created for a consumer whenever his/her *Insurance Type* or *Policy*

Number changes.

- For each reporting period, the Unknown Values Report logic captures all **Insurance_History** records having *Effective Dates* within the appropriate period.

Living_Situation table: Multiple **Living_Situation** records can appear for each consumer.

- Records in the **Living_Situation** table contain an *Effective Date*.
- For each reporting period, the Unknown Values Report logic captures all **Living_Situation** records having *Effective Dates* within the appropriate period.

SATIS_Adm table: For fields other than *Pregnancy Status* and *Week Prenatal Care Started*, the denominator is simply the number of records in the **SATIS_Adm** table having an *Admission Date* within the report period.

- For the *Pregnancy Status* field, the denominator is the number of **SATIS_Adm** records for Female Consumers.
- For the *Week Prenatal Care Started* field, the denominator is the number of **SATIS_Adm** records for female consumers whose *Pregnancy Status* field indicates that they are pregnant (NOTE: This filter was added on October 31, 2006).

SATIS_Dis table: The denominator is simply the number of records in the **SATIS_Dis** table that have *Discharge Dates* during the report period.

14. Q: Some fields have different denominators than other fields within the same table. Why does this happen?

A: The fields on the report that will have different denominators than other fields in the same table are as follows:

Discharge Type field (**Admission_History** table): The **Admission_History** table contains both *Discharge Date* and *Discharge Type* fields.

- On the report, a record's *Discharge Type* value is only included if the record has a *Discharge Date*.
- This will result in a different denominator value than that for the *Admission Type* field in the **Admission_History** table; *Admission Type* is measured for all admissions within the report period.

Pregnancy Status (**SATIS_Adm** table): The *Pregnancy Status* field in the **SATIS_Adm** table is only analyzed for female consumers. Its denominator will therefore be lower than the total number of **SATIS_Adm** records being evaluated for the report period.

Week Prenatal Care Started (**SATIS_Adm** table): The *Week Prenatal Care Started* field in the **SATIS_Adm** table is only analyzed for female consumers that are pregnant. Its denominator will therefore be lower than the denominator for the *Pregnancy Status* field in the **SATIS_Adm** table.

15. Q: We have entered a valid Language for a consumer, but we noticed that the Unknown Values Report evaluates both the Language 1 and Language 2 fields. As long as a consumer has one valid Language, we don't want the other Language field to count as missing data on the Report.
- A: For the reports issued in October 2006, new logic was added to analyze certain related fields as groups. These 'Grouped Field' analyses are as follows:

- **Client** table: *Language 1* and *Language 2* fields: If neither Language field contains a valid Language code, then both Language fields are counted as missing data.

- If one of the fields contains a valid Language code, then the other field will not be counted as missing data if it is empty.
- If the other field contains an invalid Language code, however, it will be counted as missing data.

Demographic_History table: *Address 1* and *Address 2* fields: As long as one Address field contains valid Address information or the 'UNK' code, the other field is not counted as missing data if it is empty.

SATIS_Adm and **SATIS_Dis** tables: *Axis I, II, and III Diagnosis* fields: As long as one of the nine Diagnosis fields contains a valid Diagnosis code other than 'Diagnosis Deferred' (799.9), the other fields are not counted as missing data if they are empty. Any fields containing invalid Diagnosis codes, however, are counted as missing data.

SATIS_Adm and **SATIS_Dis** tables: *Drug Type, Route of Administration, Days Used, and Age at First Use* fields: Note that the *Age at First Use* fields do not appear in the **SATIS_Dis** table.

- As long as one of the five *Drug Type* fields contains a valid *Drug Type* code other than 'None', the other *Drug Types* are not counted as missing data if they are empty or are set to 'None'.
- Note that, if any *Drug Type* field is set to an invalid *Drug Type* code, it is counted as missing data.
- For any *Drug Type* field containing a valid code other than 'None', the report logic checks the field's associated *Route of Administration, Days Used, and Age at First Use* (**SATIS_Adm** table only) fields. If any of these fields is set to empty or an invalid value, then the field is marked as missing data.

16. Q: I was told that a certain field was optional, but the Unknown Values Report is still displaying it and highlighting it if it is over the threshold percentage of Unknown values. Why is this happening?

A: As of this time, all fields are considered required.

17. Q: Who can I contact if I have questions regarding the Unknown Values Report?

A: The following DMHAS Quality Management and Improvement (QMI) Division staff act as liaisons with Providers for data quality issues:

Karin Haberlin: (860) 418-6842 E-Mail: Karin.Haberlin@po.state.ct.us

Kristen Miller: (860) 418-6749 E-Mail: Kristen.Miller@po.state.ct.us

You may contact the following staff regarding the development of Unknown Value Reports:

Jeff Johnson: (860) 418-6812 E-Mail: Jeffrey.Johnson@po.state.ct.us

Mike Hettinger: (860) 418-6858 E-Mail: Michael.Hettinger@po.state.ct.us

Jim Siemianowski is the Acting Director of the Evaluation, Quality Management, and Improvement Division. He may be reached at (860) 418-6810 or James.Siemianowski@po.state.ct.us